

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
Please do not report the images to the
Image Problem Mailbox.

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-097199

(43)Date of publication of application : 30.03.1992

(51)Int.Cl. G10L 9/14

(21)Application number : 02-209337 (71)Applicant : TOSHIBA CORP

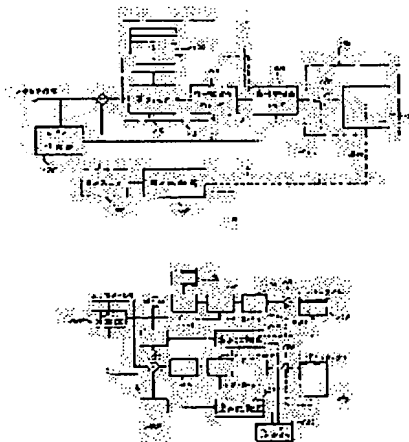
(22)Date of filing : 09.08.1990 (72)Inventor : MISEKI KIMIO
AKAMINE MASAMI

(54) VOICE ENCODING SYSTEM

(57)Abstract:

PURPOSE: To obtain a synthesized voice of high quality by providing a means stored with coefficient information on a filter and obtaining a synthesized voice signal by using the coefficient information.

CONSTITUTION: A composite filter is composed of a polarity-zero filter and the zero filter 115 has the coefficient information on the zero filter in a code book B176. Then the electric power of a weighted error signal, obtained by weighting the error signal between an input signal and a synthesized voice signal outputted by the composite filter 113 consisting of the zero filter 115, the code book B176, and a polarity filter by a weighing filter 120, is found by varying the coefficients in a code book A175 and the code book B176 in a closed loop shape. Further, when the weighted error becomes minimum, a distortion comparator 210 outputs the index of the coefficient in the code book A175 and the index of the coefficient in the code book B176 at the time of the minimum error as encoded signals corresponding to the input voice signal. Consequently, the stable synthesized voice is obtained.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑫ 公開特許公報(A) 平4-97199

⑮ Int.Cl.⁵

G 10 L 9/14

識別記号

J
H

庁内整理番号

8622-5H
8622-5H

⑬ 公開 平成4年(1992)3月30日

審査請求 未請求 請求項の数 6 (全15頁)

⑭ 発明の名称 音声符号化方式

⑯ 特 願 平2-209337

⑰ 出 願 平2(1990)8月9日

⑱ 発 明 者 三 関 公 生 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝総合
研究所内

⑲ 発 明 者 赤 嶺 政 巳 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝総合
研究所内

⑳ 出 願 人 株 式 会 社 東 芝 神奈川県川崎市幸区堀川町72番地

㉑ 代 理 人 弁 理 士 則 近 憲 佑

明 細 書

1. 発明の名称

音声符号化方式

2. 特許請求の範囲

(1) 極フィルタ及び零フィルタからなる合成フィルタを駆動信号で駆動して合成音声信号を得る音声符号化方式において、前記零フィルタの係数情報を格納する手段を有し、前記係数情報を用いて前記合成音声信号を得ることを特徴とする音声符号化方式。

(2) 極フィルタ及び零フィルタからなる合成フィルタと駆動信号で駆動して合成音声信号を得る音声符号化方式において、前記フィルタの係数情報を格納する手段を有し、前記係数情報を用いて合成音声信号を生成し、この合成音声信号と入力音声信号とのひずみにもとづいて前記零フィルタの係数情報の選択を行なうことを特徴とする音声符号化方式。

(3) 駆動信号パラメータと、極フィルタ及び零フィルタからなる合成フィルタのパラメータのビ

ット割りあてが異なる複数種類の符号化方式から各符号化方式による合成音声信号と入力音声信号のひずみを計算して1つの符号化方式を選択する音声符号化方式において、前記複数種類の符号化方式のうち少なくとも1つの符号化方式は前記零フィルタの係数情報を格納する手段を有し、前記係数情報を用いて合成音声信号を生成し、この合成音声信号と入力音声信号とのひずみにもとづいて前記零フィルタの係数の選択を行なうことを特徴とする音声符号化方式。

(4) 前記駆動信号パラメータと前記スペクトルパラメータのビット割りあてが、前記合成フィルタ中に、前記零フィルタを用いるか用いないかに依存して、決まることを特徴とする請求項2及び3記載の音声符号化方式。

(5) 前記合成フィルタ中の前記極フィルタが、各符号化方式で共通であることを特徴とする請求項2及び3記載の音声符号化方式。

(6) 前記合成フィルタのうち、前記零フィルタのフィルタ係数を入力音声信号と合成音声信号と

の聴感重み付誤差に基づいて選択することを特徴とする請求項2及び3記載の音声符号化方式。

3. 発明の詳細な説明

〔発明の目的〕

（産業上の利用分野）

この発明は音声信号等を高能率に圧縮する音声符号化方式に係り、特に低ビットの伝送レートにおける音声符号化方式に関する。

（従来の技術）

音声信号を低ビットの伝送レートで伝送する場合において、例えば10kb/s程度以下の伝送情報量で符号化する効果的な方法として、マルチモードCELP(Code Excited Linear Prediction)符号化方式が知られている。この詳細は1989年のグタスゴーで行われたICASSPの論文（第1の論文）「Multimode coding: Application to CELP Toshohiko Taniguchi, Shigeyuki Unagami and Robert M. Gray」に記載されている。この内容を簡単に説明する。第6図はそれぞれ前記論文に記載されたマルチモード符号化の原理を説明する図、

の概要である。このマルチモード符号化の考えをCELP方式に応用したものが第7図に示されるマルチモードCELP符号化器である。

CELP方式は、駆動信号のベクトル量子化を合成音のレベルで行う音声符号化方式であり、公知な技術である。又、CELP方式についての詳細は「M.R. Schroeder and B.S. Atal, "Codeexcited linear prediction CELP: High quality speech at very low bit rates.2 Proc. ICASSP'85, pp. 937-940」に記載されている。

第7図のマルチモードCELP方式は、上記のマルチモード符号化方式を2つのモードという最も簡単な形でCELPに適用したものである。すなわちAモードは、従来の公知なCELP方式で、駆動信号パラメータ、スペクトルパラメータ（LPCパラメータ）を伝送し、さらに1ビットのモード情報をフレーム毎に伝送する。

一方、Bモードはスペクトルパラメータを伝送せずに、前のフレームと同じスペクトルパラメータを用いることで、駆動信号パラメータに割りあ

第7図はマルチモードCELP符号化器の処理を示すブロック図である。

第6図において、符号側は、 m 個の符号化器510, 520, 530（符号化器51～符号化器5 m ）を備え、各符号化器は予め駆動信号パラメータとスペクトルパラメータに対して異なるビット割りあてを与えるように設定されている。

各符号化器はフレーム単位で評価と最適符号化器の決定部550で入力音声信号を並列的に処理し、入力音声信号を用いて、各符号化器の与える合成音声信号（複合音声信号）の品質を評価し、セレクタ540で最適な符号化器のインデックス n （ n は1, 2, ..., m のうちのいずれか）を用いて、伝送する駆動信号パラメータ及びスペクトルパラメータを選択し伝送すると共に、インデックス n の情報も複合側に伝送する。複合側では、符号化器のインデックス n を基に、符号化器5 n に対応する複合化器580（複合化器5 n ）を用いることにより合成音声信号を出力する。

以上が前記論文で示されたマルチモード符号化

器の量子化ビット数を増加させた構成となっている。各フレームにおいて、A/Bのモード決定は、それぞれのモードの合成音声信号の品質評価（SNR等を用いる）に基づいて行われ、伝送情報の割りあては2つのモード間のスイッチングによりダイナミックにコントロールされる。第7図において、AモードではLPC分析部100は入力音声信号からスペクトルパラメータ（LPCパラメータ）を抽出し、切り換え端子A及び短時間合成フィルタ110に出力する。長時間合成フィルタ150のパラメータ及びコードブック（小）170から選択されるベクトルの波形（コードブック内のベクトルに付されるインデックス+符号）及びゲインは入力音声と短時間合成フィルタ110（合成フィルタ）で合成された合成信号との誤差信号を、重みフィルタ120で重み付けした重み付き誤差信号の電力が最小化するように閉ループ的に求める。

一方、Bモードでは、スペクトルパラメータメモリ240がAモードと決定された場合のみ端子Aに接続されスペクトルパラメータを更新する構成

となっており、スペクトルパラメータメモリ240に蓄積されるスペクトルパラメータはBモードである間は更新されずに同じものが使用される。長時間合成フィルタ180のパラメータ及びコードブック(大)180の波形及びゲインはAモードで行ったのと同様の方法で決定される。モード決定部230はAモード、Bモードで計算された各モードの誤差電力の最小値を入力し、誤差電力の小さい方のモードを決定されたモードとして出力する。

以上が第7図のマルチモードCELP方式(従来方式)の説明である。

この方式は、従来のCELP方式に比べて4.8kbit/s及び8kbit/sの伝送レートにおいて、約2dBのセグメンタルSNRの改善があることが上記第1の論文でも示されている。

この音声符号化方式は、入力信号に応じてAモードBモードと切りかわることにより駆動信号とスペクトルパラメータのビット割り当てがフレーム毎に可変であった。

そしてフレームを一定の符号量で伝送する際、

SN比をクリアしている音声の違いはわかりにくい。

また、母音以外のスペクトルの変化の大きな音声区間はAモード(通常のCELP方式)が選択されやすいので、聴感的には通常のCELP方式による音声品質の劣化は改善されないという問題点があった。

(発明が解決しようとする課題)

上述したように、従来の音声符号化方式は、現フレームのスペクトルパラメータを使うモードと、前フレームのスペクトルパラメータを使うモードとの2つのモードの切り換えにより、駆動信号パラメータとスペクトルパラメータのビット割り当てがフレーム毎に可変であるが、スペクトルの時間的変化の大きな子音等の音声区間では前フレームのスペクトルパラメータを使用するモードは使用されにくくなるため、低レートでは結局、従来の音声符号化方式であるCELP方式における非定常区間の音声品質の劣化は改善されないという問題点がある。

本発明は、このような問題点を解決するために

Aモードではスペクトルパラメータへのビットの割りあてが多くなり、駆動信号パラメータにはあまりビットを割りあてることができない。このため、Aモードでは従来のCELP方式と同一であり、Bモードが使われる音声の区間では前のフレームと同じスペクトルパラメータを用いることにより駆動符号信号パラメータにより多くの量子化ビットを割りあてることができる。よって、BモードではCELP方式における音声品質の改善がなされる。

一方、Bモードは現フレームのスペクトルパラメータの代りに前フレームのスペクトルパラメータを使用できるような音声区間、すなわち、時間的にスペクトルの変化の少ないような母音の区間で選択されやすいことは明白である。

ところがこのような音声区間は一般に駆動信号の周期的くり返しによる冗長度も高いため、通常のCELP方式でも高いSN比の合成音声を得られる。このような音声区間にBモードの符号化を行うと、CELP方式よりもさらに高いSN比の合成音声を得られることが期待されるが聴感的にはある程度高い

なされたものであり、その目的は、低ビットの伝送レートで高品質の合成音声を得ることのできる音声符号化方式を提供することである。

[発明の構成]

(課題を解決するための手段)

上述した目的を達成するため、本発明の音声符号化方式は、極フィルタ及び零フィルタからなる合成フィルタを駆動信号で駆動して合成音声信号を得る音声符号化方式において、前記零フィルタの係数情報を格納する手段を有し、前記係数情報を用いて前記合成音声信号を得ることを特徴とするものである。

(作 用)

上述した構成を有する本発明の音声符号化方式によれば、極フィルタ及び零フィルタからなる合成フィルタのうち、該零フィルタの係数情報を格納する手段を有し、この係数情報を用いて合成音声信号を得るので、スペクトルの変化の大きな子音等の音声区間でも、該区間の音声にあったフィルタを選択することができる。よって高品質で安定

した合成音声を得ることができる。

(実施例)

以下、図面を参照して本発明の符号化方式について詳細に述べる。

第1図、第2図は本発明の音声符号化方式を行なうためのブロック図である。第1図において入力音声信号はLPC分析部100により線形予測とピッチ検出を行われ、これを短時間合成フィルタ110及び長時間合成フィルタ150に出力する。そしてコードブックA 175から選択されるベクトルの波形(該コードブックA内のベクトルに付されるインデックス+符号)及びゲインが乗算回路190を介して長時間合成フィルタ150に入力される。長時間合成フィルタ150では、入力音声信号のピッチの周期性を除去する。これを短時間合成フィルタ(以下合成フィルタという)110に入力すると、前記LPC分析部100の線形予測による予測パラメータ(合成フィルタ(極フィルタ)110の係数情報)から合成音声信号を生成する。ここで本発明によれば、合成フィルタを極零形フィル

タで構成するので、零フィルタ115を有する。そして零フィルタ115はコードブックB 176に零フィルタの係数情報を有している。よって零フィルタ115及び極フィルタからなる合成フィルタ113から出力される合成音声信号と前記入力信号との誤差信号に対して、重みフィルタ120で重み付けした重み付け誤差信号の電力を、前記コードブックA 175及びコードブックB 176内の係数を閉ループ的に変化させる。そして歪み比較器210はこれら重み付けした誤差が最小となると、該最小となる時のコードブックA 175内の係数のインデックス及びコードブックB 176内の係数のインデックスを入力音声信号に対応する符号化信号として出力する。なお、第1図の零フィルタ115に対応する第2図の $B(Z)$ が $B(Z)^{-1}$ の場合、零フィルタの係数の情報はない。ここで固定レートで伝送を行なう際、伝送できる駆動信号パラメータ及び零フィルタのパラメータは決まってしまう。しかし、一定の符号量であればこれらに対するビットの割りあては任意でもかまわない。したがって上述した

ように $B(Z)^{-1}$ の場合には零フィルタのパラメータは送る必要がなく、駆動信号パラメータにより多くのビットを割りあてることができる。反対に $B(Z)$ の場合は、零フィルタの係数も伝送しなければならないので、駆動信号パラメータのビット割りあては少なくなる。

次に第3図は第1図に示した音声符号化方式を複数用いた方式を示したブロック図である。第3図では $B(Z) \neq 1$ の場合、零フィルタ115はコードブックB 176を有しているため、零フィルタ115の $B(Z)^{-1}$ の場合における駆動信号パラメータのコードブック180より小さくなってしまふ。

さらに第4図は本発明の一実施例に係る符号化方式を符号化装置に適用した場合のブロック図を示す。

第4図において、入力端子10からA/D変換された入力音声信号の系列が入力される。フレームバッファ11は入力音声信号を1フレーム分蓄積する回路である。第4図の各ブロックはフレーム単位又はフレームを複数個に分割したサブフレーム単

位に以下の処理を行う。

予測パラメータ計算回路12は、予測パラメータを公知の方法を用いて計算する。予測フィルタが第5図に示すような長時間予測フィルタ41と短時間予測フィルタ42を縦続接続して構成される場合、予測パラメータ計算回路12はピッチ周期ピッチ予測係数および線形予測係数(α パラメータまたはKパラメータ:総してLPCパラメータと称す)を自己相関法や共分散法等の公知の方法で計算する。計算法については、例えば(古井貞昭著「デジタル音声処理」1985年東海大学出版会発行)に記載されている。計算された予測パラメータは、予測パラメータ符号化回路13へ入力される。予測パラメータ符号化回路13は、予測パラメータを予め定められた量子化ビット数に基づいて符号化し、この符号をマルチプレクサ25に出力すると共に、ゲイン計算回路15、合成フィルタ18、重みフィルタ20へそれぞれ出力する。

ゲイン計算回路15は後述する零フィルタ係数コードブック14からの零フィルタの係数と、係数検

索回路24から出力される係数更新信号と、符号化回路13からの予測パラメータ（極フィルタの係数情報）をもとに極零形の合成フィルタ $H(Z)$ を構成する。この逆フィルタ $1/H(Z)$ を予測フィルタとして入力音声信号を予測し、予測残差信号を作成する。次にゲイン計算回路15は予測残差信号の平均パワーを計算してこれをゲインとして符号化回路16へ出力する。前記予測残差信号の平均パワーとしては、例えば標準偏差を用いることができる。符号化回路16はゲインを予め定められた量子化ビット数に基づいて符号化し、この符号をマルチプレクサ25および乗算回路17へ出力する。零フィルタ係数コードブック14は予め定められた次数と、量子化ビット数 M に対応した $2M$ 種類の零フィルタのフィルタ係数情報を格納するものである。また、零フィルタ係数コードブック14に格納される零フィルタ $B(Z)$ の1つに $B(Z)-1$ となるフィルタ情報を格納すれば、零フィルタを用いない全極形の合成フィルタ $H(Z)$ が自動的に同一の構成で作成できる。

本実施例では、零フィルタ係数コードブック14

は無いので、その分駆動信号に多くのビット数割りあてて、駆動信号の形状を表すコードブック21内のコードベクトルの検索範囲を広げることができる。

逆に、該情報 PZ が零フィルタ $B(Z) \times 1$ を示す情報である場合は零フィルタ係数の情報を伝送する必要があるので、その分駆動信号に少ないビット数を割りあてて、コードブック21内のコードベクトルの検索範囲をせばめるものとする。

乗算回路17は、コードブック21から出力されるコードベクトルに符号化されたゲインを乗じて駆動信号の候補となるベクトルを生成し、合成フィルタ18へ入力する。

合成フィルタ18は零フィルタ係数コードブック14と符号化回路13とより、零フィルタの係数情報および極フィルタの係数情報（これをまとめてスベクトルパラメータと呼んでいる）をそれぞれ入力し、合成フィルタ $H(Z)$ を構成し、乗算回路17よりの駆動信号の候補ベクトルを入力信号として合成音声信号を出力する。

は、 $2M+1$ 種類の零フィルタ係数情報を格納し、その第1番のコードベクトルを用いて作成される零フィルタ $B(Z)$ は、 $B(Z)-1$ となるように予めコードブック14が作成されているものとする。

零フィルタ係数コードブック14は、係数探索回路24から入力されるコード更新信号に基づき、該零フィルタコードブック14に格納された零フィルタ係数（コードベクトル）をゲイン計算回路15、合成フィルタ18へ出力すると共に、零フィルタ $B(Z)$ が $B(Z)-1$ か $B(Z) \times 1$ かの情報 PZ をコードブック21へ出力する。

コードブック21はコードブック14からの情報 PZ に応じて予め設定される制限された数のコードベクトルを乗算回路17へ出力する。このときのコードベクトルの出力は、コード探索回路23から入力されるコード更新信号によって制御される。コードブック21内のコードベクトルの検索範囲の制限は例えば次のように決めることができる。

コードブックからの情報 PZ が零フィルタ $B(Z)-1$ を示す情報である場合は、零フィルタ係数の情報

乗算回路19は入力音声信号と上述の合成音声信号を入力し、その誤差信号を出力する。

重みフィルタ20は上述の誤差信号に予測パラメータから作成される重みを付けて出力する。重みフィルタ20は伝達関数が

$$V(Z) = \frac{A(Z)}{A(Z/\tau)} \quad (0 \leq \tau \leq 1) \quad (1)$$

で表されるフィルタで、聴覚のマスクング効果を利用して、複合時に合成音声に含まれる符号化ノイズを聞こえにくくする効果があることが知られている。(1)式において、 $A(Z)$ は予測パラメータから作成される予測フィルタを表している。

2乗誤差計算回路22は、重み付けされた誤差信号の2乗和をコードブック21から出力されるコードベクトル毎に計算し、その結果をコード探索回路23へ出力すると共に、誤差信号の2乗和を1フレーム分計算した値を係数探索回路24へ出力する。

コード探索回路23は後述する係数探索回路24から出力される現在検索中の零フィルタのコード番号を入力し、その零フィルタのコード番号ごとに

各サブフレームの2乗誤差が最小となるコードをコードブック21から検索し、このコードを保持する。係数検索回路24で最終的に零フィルタのコード番号が決定すると、この番号を入力し保持していた駆動信号のコードのうち、零フィルタのコード番号に対応して保持していたコードをマルチプレクサ25へ出力する。

係数検索回路24は2乗誤差計算回路22から入力される各零フィルタのコード番号毎にフレーム単位で計算された誤差信号の2乗和を比較してこれが最小となる零フィルタのコード番号を選択し、このコード番号をマルチプレクサ25およびコード検索回路23へ出力する。もし検索された零フィルタ係数のコード番号が1ならば上述したように、零フィルタは非使用であることがわかるので、このときは、コード検索回路23から出力される駆動信号のコードは零フィルタ使用時に比べてより大きなビット数で表されている。係数検索回路24は零フィルタの使用・非使用の情報も同時にマルチプレクサ25へ出力する。第1表に本実施例におけ

る駆動信号とスペクトルパラメータとの間のビット配分の例を示す。

る駆動信号とスペクトルパラメータとの間のビット配分の例を示す。

使用する合成フィルタ	スペクトルパラメータ用ビット数	駆動信号用ビット数	フレームあたりのビット数
全極フィルタ	K	R-K	R
極零フィルタ	K+M	R-K-M	R

第1表

第1表において、使用する合成フィルタは零フィルタが $B(Z)=1$ と $B(Z)=1$ の場合により、全極フィルタと極零フィルタとに分けることができる。今、フレームあたりのビット数をRビットとする時、スペクトルパラメータ用ビット数は極フィルタのビット数Kビットのみとなり、駆動信号要ビット数は当然 $R-K$ ビットとなる。よってフレームあたりのビット数は常にR一定となる。また、極零フィルタを用いた場合には、零フィルタにもスペクトルパラメータ用ビットとしてMビットを割りふるので、残りを駆動用信号とするものである。マルチプレクサ25は入力されるコード情報を多

重化し、端子26より伝送路へコード情報を出力する。

このように、本発明の音声符号化によれば、入力音声信号の音質の変化に反応して、スペクトル包絡を表すフィルタと駆動信号のパラメータのビット配分がフレーム単位で変化するだけでなく、このフィルタを極零形で表し、零フィルタのフィルタ係数の量子化、つまりコードブックの選択を、入力音声信号と合成音声信号の聴感重み付けした誤差が最小となるように行っている。このため、スペクトルお時間的変化が大きな音声区間に対しても、その区間に適合したフィルタを選択できるので、合成音声の品質を安定して向上させることができる。

なお、ここで説明した実施例は本発明の一実施例であり、様々な変形が可能である。

【効果】

以上詳述したように本発明の音声符号化方式によれば、高品質で安定した合成音声を得ることができる。

4. 図面の簡単な説明

第1図、第2図は本発明の音声符号化方式を行うためブロック図、第3図は複数の音声符号化方式に本発明の音声符号化方式を用いたブロック図、第4図は本発明の一実施例に係る音声符号化方式を符号化装置に適用した構成を示すブロック図、第5図は第4図を用いた実施例に記載される予測フィルタの一構成例を示すブロック図、第6図、第7図は従来技術による符号化装置の構成を示すブロック図である。

110 … 短時間合成フィルタ (極フィルタ)

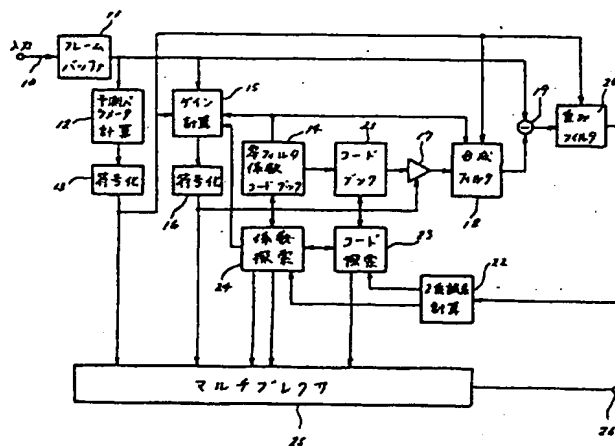
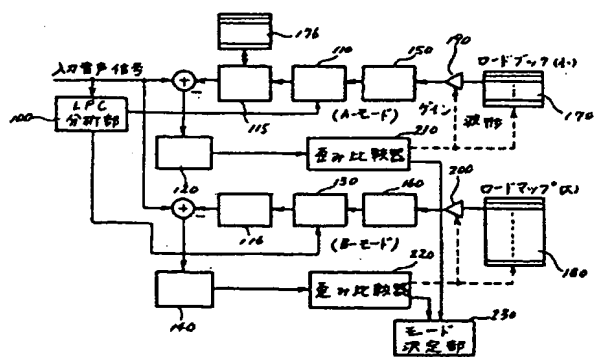
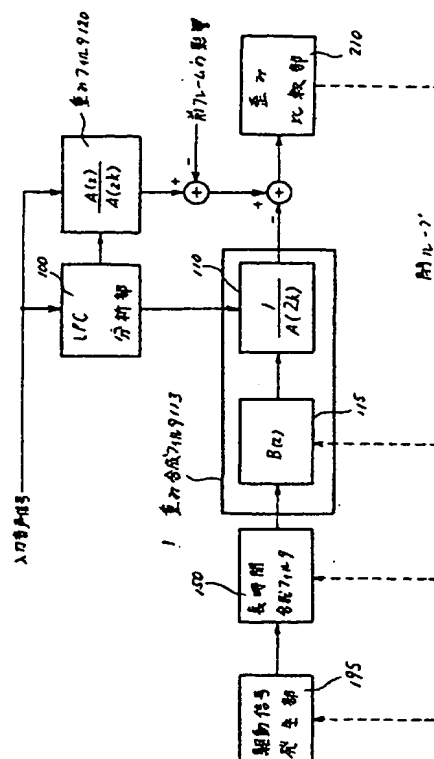
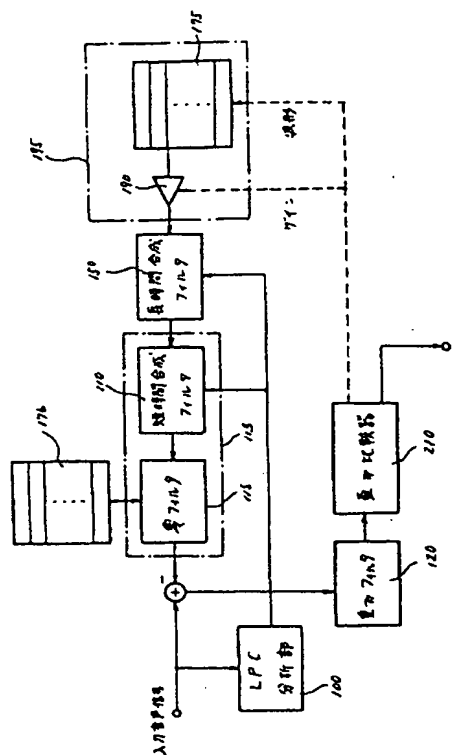
113 … 合成フィルタ

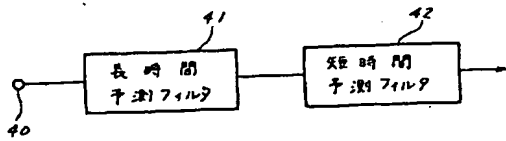
115 … 零フィルタ

175, 176 … コードブック

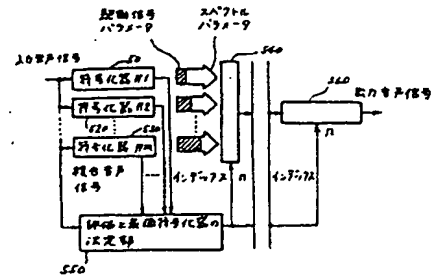
195 … 駆動信号発生部

代理人 弁理士 則 近 憲 佑

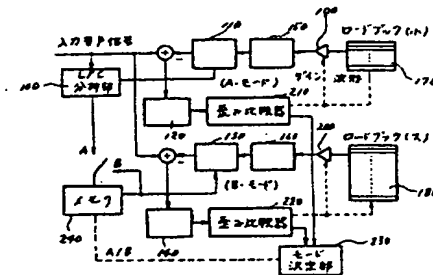




第 5 図



第 6 図



第 7 図

手続補正書 (自発)

2.10.12

特許庁長官 殿 平成 年 月 日

1. 事件の表示

平成 2 年特許願第 209337 号

2. 発明の名称

音声符号化方式

3. 補正をする者

事件との関係 特許出願人

(307) 株式会社 東芝

4. 代理人

〒105 東京都港区芝浦一丁目1番1号
株式会社東芝 本社事務所内
電話 (03)457-2512 (代表)

(7317) 弁護士 則近 幸佑

5. 補正の対象

- (1) 図面
- (2) 明細書全文

6. 補正の内容

(1) 明細書の全文の記載を別紙のとおり補正する。

(2) 図面の第 1 図乃至第 4 図及び第 6 図乃至第 7 図を別紙のとおり補正する。

以上



訂正明細書

1. 発明の名称

音声符号化方式

2. 特許請求の範囲

(1) 極フィルタ及び零フィルタからなる合成フィルタを駆動信号で駆動して合成音声信号を得る音声符号化方式において、前記零フィルタの係数情報を格納する手段を有し、前記係数情報を用いて前記合成音声信号を得ることを特徴とする音声符号化方式。

(2) 極フィルタ及び零フィルタからなる合成フィルタと駆動信号で駆動して合成音声信号を得る音声符号化方式において、前記フィルタの係数情報を格納する手段を有し、前記係数情報を用いて合成音声信号を生成し、この合成音声信号と入力音声信号とのひずみにもとづいて前記零フィルタの係数情報の選択を行なうことを特徴とする音声符号化方式。

(3) 駆動信号パラメータと、極フィルタ及び零フィルタからなる合成フィルタのパラメータの

(産業上の利用分野)

この発明は音声信号等を高効率に圧縮する音声符号化方式に係り、特に低ビットの伝送レートにおける音声符号化方式に関する。

(従来技術)

音声信号を低ビットの伝送レートで伝送する場合において、例えば10 kb/s程度以下の伝送レートで符号化する効果的な方法として、マルチモードCELP (Code Excited Linear Prediction) 符号化方式が知られている。この詳細は1989年のグラスゴーで行われたICASSPの論文(第1の論文)「Multimode coding: Application to CELP Tomohiko Taniguchi, Shigeyuki Unagami and Robert M. Gray」に記載されている。この内容を簡単に説明する。第6図はそれぞれ前記論文に記載されたマルチモード符号化の原理を説明する図、第7図はマルチモードCELP符号化器の処理を示すブロック図である。

第6図において、符号側は、 m 個の符号化器510、520、530(符号化器#1～符号化

器# m)を備え、各符号化器は予め駆動信号パラメータとスペクトルパラメータに対して異なるビット割りあてを与えるように設定されている。各符号化器はフレーム単位で評価と最適符号化器の決定部550で入力音声信号を並列的に処理し、入力音声信号を用いて、各符号化器の与える合成音声信号(復号音声信号)の品質を評価し、セレクト540で最適な符号化器のインデックス n (n は1、2、... m のうちいずれか)を用いて、伝送する駆動信号パラメータ及びスペクトルパラメータを選択し伝送すると共に、インデックス n の情報も復号側に伝送する。復号側では、符号化器のインデックス n を基に、符号化器# n に対応する復号化器560(復号化器# n)を用いることにより合成音声信号を出力する。

(4) 前記駆動信号パラメータと前記合成フィルタのパラメータのビット割りあてが、前記合成フィルタ中に、前記零フィルタを用いるか用いなくに依存して、決まることを特徴とする請求項2及び3記載の音声符号化方式。

(5) 前記合成フィルタ中の前記極フィルタが、各符号化方式で共通であることを特徴とする請求項2及び3記載の音声符号化方式。

3. 発明の詳細な説明

[発明の目的]

器# m)を備え、各符号化器は予め駆動信号パラメータとスペクトルパラメータに対して異なるビット割りあてを与えるように設定されている。

各符号化器はフレーム単位で評価と最適符号化器の決定部550で入力音声信号を並列的に処理し、入力音声信号を用いて、各符号化器の与える合成音声信号(復号音声信号)の品質を評価し、セレクト540で最適な符号化器のインデックス n (n は1、2、... m のうちいずれか)を用いて、伝送する駆動信号パラメータ及びスペクトルパラメータを選択し伝送すると共に、インデックス n の情報も復号側に伝送する。復号側では、符号化器のインデックス n を基に、符号化器# n に対応する復号化器560(復号化器# n)を用いることにより合成音声信号を出力する。

以上が前記論文で示されたマルチモード符号化の概要である。このマルチモード符号化の考えをCELP方式に応用したものが第7図に示されるマルチモードCELP符号化器である。

CELP方式は、駆動信号のベクトル量子化を

合成音のレベルで行う音声符号化方式であり、公知な技術である。又、CELP方式についての詳細は「M.R. Schroeder and B.S. Atal, "Code-excited linear prediction CELP: High quality speech at very low bit rates, 2 Proc. ICASSP '85, pp. 937-940」に記載されている。

第7図にマルチモードCELP方式は、上記のマルチモード符号化方式を2つのモードという最も簡単な形でCELPに適用したものである。すなわちAモードは、従来の公知なCELP方式で、駆動信号パラメータ、スペクトルパラメータ(LPCパラメータ)を伝送し、さらに1ビットのモード情報をフレーム毎に伝送する。

一方、Bモードはスペクトルパラメータを伝送せずに、前のフレームと同じスペクトルパラメータを用いることで、駆動信号パラメータに割りあてられる量子化ビット数を増加させた構成となっている。各フレームにおいて、A/Bのモード決定は、それぞれのモードの合成音声信号の品質評価(SNR等を用いる)に基づいて行われ、伝送情

報の割りあては2つのモード間のスイッチングによりダイナミックにコントロールされる。第7図において、AモードではLPC分析部100は入力音声信号からスペクトルパラメータ(LPCパラメータ)を抽出し、切り換え端子A及び短時間合成フィルタ110に出力する。長時間合成フィルタ150のパラメータ及びコードブック(小)170から選択されるベクトルの波形(コードブック内のベクトルに付されるインデックス+符号)及びゲインは入力音声と短時間合成フィルタ110(合成フィルタ)で合成された合成信号との誤差信号を、重みフィルタ120で重み付けした重み付き誤差信号の電力が最小化するように閉ループ的に求める。

一方、Bモードでは、スペクトルパラメータメモリ240がAモードと決定された場合のみ端子Aに接続されスペクトルパラメータを更新する構成となっており、スペクトルパラメータメモリ240に蓄積されるスペクトルパラメータはBモードである間は更新されずに同じものが使用さ

れる。長時間合成フィルタ160のパラメータ及びコードブック(大)180の波形及びゲインはAモードで行ったのと同様の方法で決定される。モード決定部230はAモード、Bモードで計算された各モードの誤差電力の最小値を入力し、誤差電力の小さい方のモードを決定されたモードとして出力する。

以上が第7図のマルチモードCELP方式(従来方式)の説明である。

この方式は、従来のCELP方式に比べて4.8 kbit/s呼び8 kbit/sの伝送レートにおいて、約2 dBのセグメンタルSNRの改善があることが上記第1の論文でも示されている。

この音声符号化方式は、入力信号に応じてAモードBモードと切りかわることにより駆動信号とスペクトルパラメータのビット割り当てがフレーム毎に可変であった。

そしてフレームを一定の符号量で伝送する際、Aモードではスペクトルパラメータへのビットの割りあてが多くなり、駆動信号パラメータにはあ

まりビットを割りあてることができない。このため、Aモードでは従来のCELP方式と同一であり、Bモードが使われる音声の区間では前のフレームと同じスペクトルパラメータを用いることにより駆動符号信号パラメータにより多くの量子化ビットを割りあてることができる。よって、BモードではCELP方式における音声品質の改善がなされる。

一方、Bモードでは現フレームのスペクトルパラメータの代りに前フレームのスペクトルパラメータを使用できるような音声区間、すなわち、時間的にスペクトルの変化の少ないような母音の区間で選択されやすいことは明白である。

ところがこのような音声区間は一般に駆動信号の周期的くり返しによる冗長度も高いため、通常のCELP方式でも高いSN比の合成音声を得られる。このような音声区間にBモードの符号化を行うと、CELP方式よりもさらに高いSN比の合成音声を得られることが期待されるが聴感的にはある程度高いSN比をクリアしている音声の逆

いはわかりにくい。

また、母音以外のスペクトルの変化の大きな音声区間はAモード（通常のC E L P方式）が選択されやすいので、聴感的には通常のC E L P方式による音声品質の劣化は改善されないという問題点があった。

（発明が解決しようとする課題）

上述したように、従来の音声符号化方式は、現フレームのスペクトルパラメータを使うモードと、前フレームのスペクトルパラメータを使うモードとの2つのモードの切り換えにより、駆動信号パラメータとスペクトルパラメータのビット割りあてがフレーム毎に可変であるが、スペクトルの時間的変化の大きな子音等の音声区間では前フレームのスペクトルパラメータを使用するモードは使用されにくくなるため、低レートでは結局、従来の音声符号化方式であるC E L P方式における非定常区間の音声品質の劣化は改善されないという問題点がある。

本発明は、このような問題点を解決するために

で安定した合成音声を得ることができる。

（実施例）

以下、図面を参照して本発明の符号化方式について詳細に述べる。

第1図、第2図は本発明の音声符号化方式を行なうためのブロック図である。第1図において入力音声信号はL P C分析部100により線形予測とピッチ抽出が行なわれ、これによって得られるフィルタのパラメータを短時間合成フィルタ110及び長時間合成フィルタ150に出力する。そしてコードブックA175から選択されるベクトルの波形（該コードブックA内のベクトルに付されるインデックス+符号）及びゲインが乗算回路190を介して長時間合成フィルタ150に入力される。長時間合成フィルタ150は、入力信号にピッチの周期性を付加する。これを短時間合成フィルタ（以下合成フィルタという）110に入力すると、前記L P C分析部100の線形予測による予測パラメータ（合成フィルタ（極フィルタ）110の係数情報）から合成音声信号を生成

なされたものであり、その目的は、低ビットの伝送レートで高品質の合成音声を得ることのできる音声符号化方式を提供することである。

〔発明の構成〕

（課題を解決するための手段）

上述した目的を達成するため、本発明の音声符号化方式は、極フィルタ及び零フィルタからなる合成フィルタを駆動信号で駆動して合成音声信号を得る音声符号化方式において、前記零フィルタの係数情報を格納する手段を有し、前記係数情報を用いて前記合成音声信号を得ることを特徴とするものである。

（作用）

上述した構成を有する本発明の音声符号化方式によれば、極フィルタ及び零フィルタからなる合成フィルタのうち、該零フィルタの係数情報を格納する手段を有し、この係数情報を用いて合成音声信号を得るので、スペクトルの変化の大きな子音等の音声区間でも、該区間の音声にあったフィルタを選択することができる。よって高品質

する。ここで本発明によれば、合成フィルタを極零形フィルタで構成するので、零フィルタ115を有する。そして零フィルタ115はコードブックB176に零フィルタの係数情報を有している。よって零フィルタ115、コードブックB176及び極フィルタからなる合成フィルタ113から出力される合成音声信号と前記入力信号との誤差信号に対して、重みフィルタ120で重み付けした重み付け誤差信号の電力を、前記コードブックA175及びコードブックB176内の係数を閉ループ的に変化させて求める。そして重み比較器210はこれら重み付けした誤差が最小となると、該最小となる時のコードブックA175内の係数のインデックス及びコードブックB176内の係数のインデックスを入力音声信号に対応する符号化信号として出力する。第2図は入力信号と合成音声信号との誤差を重み付けして重みを評価する構成にした本発明のブロック図である。第2図において長時間合成フィルタ150のパラメータは閉ループ的に求められるように構成されている。

従ってこの場合、LPC分析部100では短時間合成フィルタのパラメータだけが求められる。なお、第1図の零フィルタ115に対応する第2図の $B(Z)$ が $B(Z) = 1$ の場合、零フィルタの係数の情報はない。ここで固定レートで伝送を行なう際、1フレームに割りあてられるビット数は決まってしまう。しかし、一定の符号量であれば各パラメータに対するビットの割りあては任意でもかまわない。したがって上述したように $B(Z) = 1$ の場合には零フィルタのパラメータは送る必要がなく、駆動信号パラメータにより多くのビットを割りあてることができる。反対に $B(Z) \neq 1$ の場合は、零フィルタの係数も伝送しなければならないので、駆動信号パラメータのビット割りあてを少なくすることで伝送レートを一定にすることができるとなる。

次に第3図は第1図に示した音声符号化方式を複数用いた方式を示したブロック図である。第3図では $B(Z) \neq 1$ の場合、零フィルタ115はコードブックB176を有しているため、駆動信

号パラメータと称す)を自己相関法や共分散法等の公知の方法で計算する。計算法については、例えば(古井貞昭著「ディジタル音声処理」1985年東海大学出版会発行)に記述されている。計算された予測パラメータは、予測パラメータ符号化回路13へ入力される。予測パラメータ符号化回路13は、予測パラメータを予め定められた量子化ビット数に基づいて符号化し、この符号化をマルチプレクサ25に出力すると共に、ゲイン計算回路15、合成フィルタ18、重みフィルタ20へそれぞれ出力する。

ゲイン計算回路15は後述する零フィルタ係数コードブック14からの零フィルタの係数と、係数検索回路24から出力される係数更新信号と、符号化回路13からの予測パラメータ(極フィルタの係数情報)をもとに極零形の合成フィルタ $H(Z)$ を構成する。この逆フィルタ $1/H(Z)$ を予測フィルタとして入力音声信号を予測し、予測残差信号を作成する。次にゲイン計算回路15は予測残差信号の平均パワーを計算してこれをゲ

インとして符号化回路16へ出力する。前記予測残差信号の平均パワーとしては、例えば標準偏差を用いることができる。符号化回路16はゲインを予め定められた量子化ビット数に基づいて符号化し、この符号化をマルチプレクサ25および乗算回路17へ出力する。零フィルタ係数コードブック14は予め定められた次数と、量子化ビット数 M に対応した種類数の零フィルタのフィルタ係数情報を格納するものである。また、零フィルタ係数コードブック14に格納される零フィルタ $B(Z)$ の1つに $B(Z) = 1$ となるフィルタ情報を格納すれば、零フィルタを用いない全極形の合成フィルタ $H(Z)$ が自動的に同一の構成で作成できる。

さらに第4図は本発明の一実施例に係る符号化方式を符号化装置に適用した場合のブロック図を示す。

第4図において、入力端子10からA/D変換された入力音声信号の系列が入力される。フレームバッファ11は入力音声信号を1フレーム分蓄積する回路である。第4図の各ブロックはフレーム単位又はフレームを複数個に分割したサブフレーム単位に以下の処理を行う。

予測パラメータ計算回路12は、予測パラメータを公知の方法を用いて計算する。予測フィルタが第5図に示すような長時間予測フィルタ41と短時間予測フィルタ42を縦続接続して構成される場合、予測パラメータ計算回路12はビッチ周期ビッチ予測係数および線形予測係数(α パラメータまたは K パラメータ:総してLPCパラ

メータと称す)を自己相関法や共分散法等の公知の方法で計算する。計算法については、例えば(古井貞昭著「ディジタル音声処理」1985年東海大学出版会発行)に記述されている。計算された予測パラメータは、予測パラメータ符号化回路13へ入力される。予測パラメータ符号化回路13は、予測パラメータを予め定められた量子化ビット数に基づいて符号化し、この符号化をマルチプレクサ25に出力すると共に、ゲイン計算回路15、合成フィルタ18、重みフィルタ20へそれぞれ出力する。

ゲイン計算回路15は後述する零フィルタ係数コードブック14からの零フィルタの係数と、係数検索回路24から出力される係数更新信号と、符号化回路13からの予測パラメータ(極フィルタの係数情報)をもとに極零形の合成フィルタ $H(Z)$ を構成する。この逆フィルタ $1/H(Z)$ を予測フィルタとして入力音声信号を予測し、予測残差信号を作成する。次にゲイン計算回路15は予測残差信号の平均パワーを計算してこれをゲ

零フィルタ係数コードブック14は、係数探索回路24から入力されるコード更新信号に基づき、該零フィルタコードブック14に格納された零フィルタ係数(コードベクトル)をゲイン計算回路15、合成フィルタ18へ出力すると共に、零フィルタB(Z)がB(Z)-1かB(Z)≠1かの情報PZをコードブック21へ出力する。

コードブック21は予め分散値が正規化されており、コードブック14からの情報PZに応じて予め設定される制限された数のコードベクトルを乗算回路17へ出力する。このときのコードベクトルの出力は、コード探索回路23から入力されるコード更新信号によって制御される。コードブック21内のコードベクトルの検索範囲の制限は例えば次のように決めることができる。

コードブックからの情報PZが零フィルタB(Z)-1を示す情報である場合は、零フィルタ係数の情報は無いので、その分駆動信号に多くのビット数割りあてて、駆動信号の形状を表すコードブック21内のコードベクトルの検索範囲を広

げることができる。

逆に、該情報PZが零フィルタB(Z)≠1を示す情報である場合は零フィルタ係数の情報を伝送する必要があるので、その分駆動信号に少ないビット数を割りあてて、コードブック21内のコードベクトルの検索範囲をせばめるものとする。

乗算回路17は、コードブック21から出力されるコードベクトルに符号化されたゲインを乗じて駆動信号の候補となるベクトルを生成し、合成フィルタ18へ入力する。

合成フィルタ18は零フィルタ係数コードブック14と符号化回路13とより、零フィルタの係数情報および極フィルタの係数情報(これをまとめてスペクトルパラメータと呼んでいる)をそれぞれ入力し、合成フィルタH(Z)を構成し、乗算回路17よりの駆動信号の候補ベクトルを入力信号として合成音声信号を出力する。

減算回路19は入力音声信号と上述の合成音声信号を入力し、その誤差信号を出力する。

重みフィルタ20は上述の誤差信号に予測パラ

メータから作成される重みを付けて出力する。重みフィルタ20は伝達関数が

$$W(Z) = \frac{A(Z)}{A(Z/\gamma)} \quad (0 \leq \gamma \leq 1) \quad (1)$$

で表されるフィルタで、聴覚のマスクング効果を利用して復号時に合成音声に含まれる符号化ノイズを聞こえにくくする効果があることが知られている。(1)式において、A(Z)は予測パラメータから作成される予測フィルタを表している。

2乗誤差計算回路22は、重み付けされた誤差信号の2乗和をコードブック21から出力されるコードベクトル毎に計算し、その結果をコード検索回路23へ出力すると共に、誤差信号の2乗和を1フレーム分計算した値を係数検索回路24へ出力する。

コード検索回路23は後述する係数検索回路24から出力される現在検索中の零フィルタのコード番号を入力し、その零フィルタのコード番号ごとに各サブフレームの2乗誤差が最小となるコードをコードブック21から検索し、このコードを保持する。係数検索回路24で最終的に零

フィルタのコード番号が決定すると、この番号を入力し保持していた駆動信号のコードのうち、零フィルタのコード番号に対応して保持しているコードをマルチプレクサ25へ出力する。

係数検索回路24は2乗誤差計算回路22から入力される各零フィルタのコード番号毎にフレーム単位で計算された誤差信号の2乗和を比較してこれが最小となる零フィルタのコード番号を選択し、このコード番号をマルチプレクサ25およびコード検索回路23へ出力する。もし検索された零フィルタ係数のコード番号が1ならば上述したように、零フィルタは非使用であることがわかるので、このときは、コード検索回路23から出力される駆動信号のコードは零フィルタ使用時に比べてより大きなビット数で表されている。係数検索回路24は零フィルタの使用・非使用の情報も同時にマルチプレクサ25へ出力する。第1表に本実施例における駆動信号とスペクトルパラメータとの間のビット配分の例を示す。

使用する 合成フィルタ	スペクトルパラメータ 用ビット数	駆動信号 用ビット数	フレームあたりの ビット数
全極フィルタ	K	R-K	R
極零フィルタ	K+M	R-K-M	R

第 1 表

第 1 表において、使用する合成フィルタは零フィルタが $B(Z) = 1$ と $B(Z) \neq 1$ の場合により、全極フィルタと極零フィルタとに分けることができる。今、フレームあたりのビット数を R ビットとする時、スペクトルパラメータ用ビット数は極フィルタのビット数 K ビットのみとなり、駆動信号用ビット数は当然 $R - K$ ビットとなる。よってフレームあたりのビット数は常に R 一定となる。また、極零フィルタを用いた場合には、零フィルタにもベクトルパラメータ用ビットとして M ビットを割りふるので、残りを駆動用信号とするものである。マルチプレクサ 25 は入力されるコード情報を多重化し、端子 26 より伝送路へコード情報を出力する。

なうためブロック図、第 3 図は複数の音声符号化方式に本発明の音声符号化方式を用いたブロック図、第 4 図は本発明の一実施例に係る音声符号化方式を符号化装置に適用した構成を示すブロック図、第 5 図は第 4 図を用いた実施例に記載される予測フィルタの一構成例を示すブロック図、第 6 図、第 7 図は従来技術による符号化装置の構成を示すブロック図である。

- 110 … 短時間合成フィルタ（極フィルタ）
- 113 … 合成フィルタ
- 115 … 零フィルタ
- 175, 176 … コードブック
- 195 … 駆動信号発生部

代理人弁理士 則 近 意 祐

このように、本発明の音声符号化によれば、入力音声信号の音質の変化に適応して、スペクトル包絡を表すフィルタと駆動信号のパラメータのビット配分がフレーム単位で変化するだけでなく、このフィルタを極零形で表し、零フィルタのフィルタ係数の量子化、つまりコードブックの選択を、入力音声信号と合成音声信号の聴感重み付けした誤差が最小となるように行っている。このため、スペクトルの時間的変化が大きな音声区間に対しても、その区間に適合したフィルタを選択できるので、合成音声の品質を安定して向上させることができる。

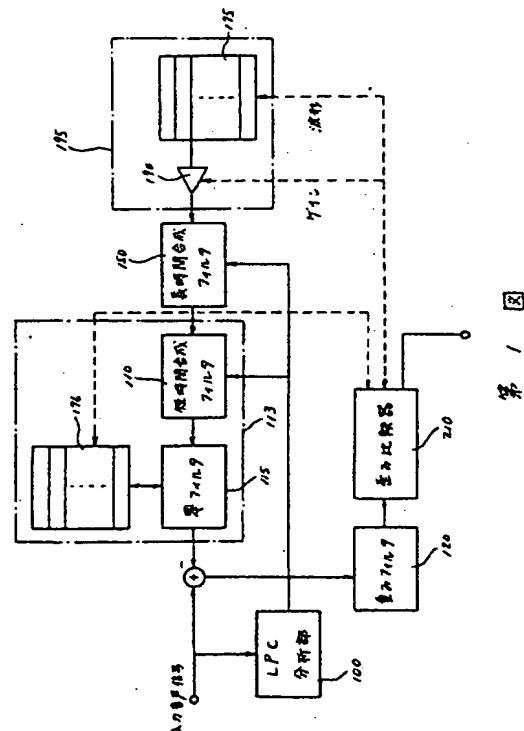
なお、ここで説明した実施例は本発明の一実施例であり、様々な変形が可能である。

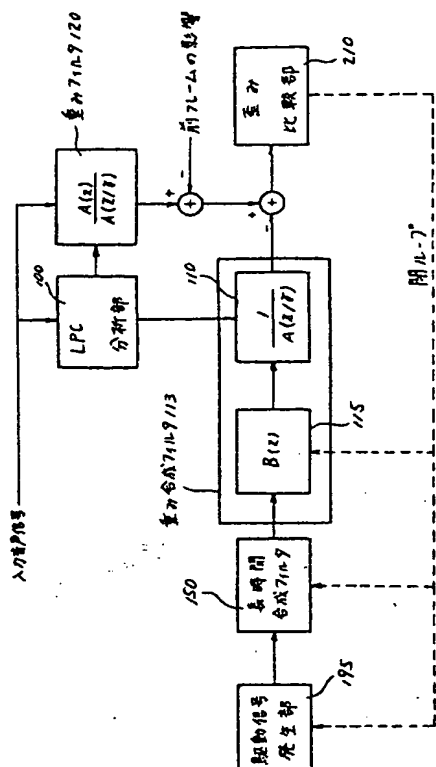
（発明の効果）

以上詳述したように本発明の音声符号化方式によれば、高品質で安定した合成音声を得ることができる。

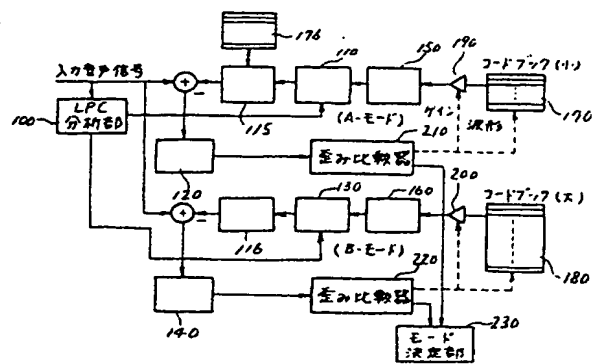
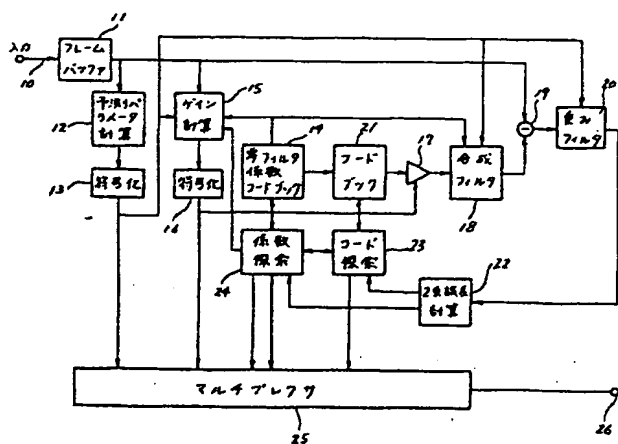
4. 図面の簡単な説明

第 1 図、第 2 図は本発明の音声符号化方式を行

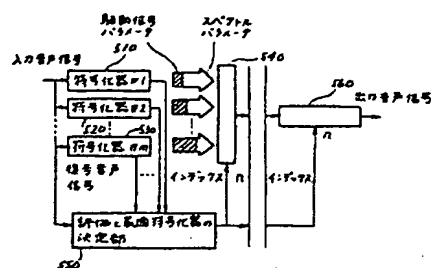




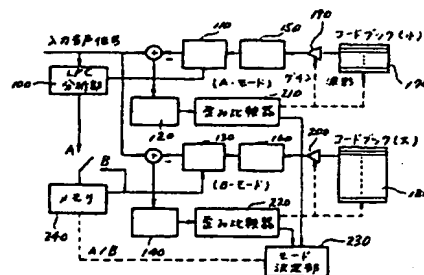
第二章

第 3 ☒

第 4 圖



第 6 回



第 7 回